Title
A snoRNA modulates mRNA 3' end processing and regulates the expression of a subset of mRNAs.

Journal
Nucleic acids research, 45(15)

ISSN
0305-1048

Authors
Huang, Chunliu
Shi, Junjie
Guo, Yibin
et al.

Publication Date
2017-09-01

DOI
10.1093/nar/gkx651

Peer reviewed

Published online 26 July 2017

Nucleic Acids Research, 2017, Vol. 45, No. 15 ***8647–8660***
doi: 10.1093/nar/gkx651

NAR Breakthrough Article

A snoRNA modulates mRNA 3′ end processing and regulates the expression of a subset of mRNAs

Chunliu Huang[1,†], Junjie Shi[1,†], Yibin Guo[2,†], Weijun Huang[1,2], Shanshan Huang[1], Siqi Ming[3], Xingui Wu[1], Rui Zhang[4], Junjun Ding[1,5], Wei Zhao[1], Jie Jia[5], Xi Huang[3], Andy Peng Xiang[1], Yongsheng Shi[6,*] and Chengguo Yao[1,5,*]

[1]Center for Stem Cell Biology and Tissue Engineering, Key Laboratory for Stem Cells and Tissue Engineering, Ministry of Education, Sun Yat-Sen University, Guangzhou 510080, China, [2]Department of Medical Genetics, Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou 510080, China, [3]Institute of Tuberculosis Control, Key laboratory of Tropical Diseases Control, Ministry of Education, Sun Yat-sen University, Guangzhou 510080, China, [4]Key Laboratory of Gene Engineering of the Ministry of Education, State Key Laboratory of Biocontrol, School of Life Sciences, Sun Yat-Sen University, Guangzhou 510275, China, [5]Department of Biology, Zhongshan School of Medicine, Sun Yat-Sen University, Guangzhou 510080, China and [6]Department of Microbiology and Molecular Genetics, School of Medicine, University of California Irvine, Irvine, CA 92697, USA

Received March 23, 2017; Revised July 12, 2017; Editorial Decision July 13, 2017; Accepted July 15, 2017

ABSTRACT

mRNA 3′ end processing is an essential step in gene expression. It is well established that canonical eukaryotic pre-mRNA 3′ processing is carried out within a macromolecular machinery consisting of dozens of trans-acting proteins. However, it is unknown whether RNAs play any role in this process. Unexpectedly, we found that a subset of small nucleolar RNAs (snoRNAs) are associated with the mammalian mRNA 3′ processing complex. These snoRNAs primarily interact with Fip1, a component of cleavage and polyadenylation specificity factor (CPSF). We have functionally characterized one of these snoRNAs and our results demonstrated that the U/A-rich SNORD50A inhibits mRNA 3′ processing by blocking the Fip1-poly(A) site (PAS) interaction. Consistently, SNORD50A depletion altered the Fip1–RNA interaction landscape and changed the alternative polyadenylation (APA) profiles and/or transcript levels of a subset of genes. Taken together, our data revealed a novel function for snoRNAs and provided the first evidence that non-coding RNAs may play an important role in regulating mRNA 3′ processing.

INTRODUCTION

All metazoan pre-mRNAs require extensive processing during their maturation. One essential step is 3′ end processing (1–4). Canonical mRNA 3′ processing involves an endonucleolytic cleavage within the pre-mRNA sequences and the addition of a poly(A) tail to the 3′ end of the transcript, and both steps take place within a macromolecular machinery called mRNA 3′ processing complex (5–7). There are ∼85 trans-acting proteins in the human mRNA 3′ processing complex, including the poly(A) polymerase (PAP) and four core multi-subunit complexes (CPSF, CstF, CF I and CF II), as well as other peripheral factors that associate with other biological processes (7). Despite much progress in deciphering the protein–protein and RNA–protein interaction within the mRNA 3′ processing complex (8–12), it remains poorly understood how a specific poly(A) site (PAS) is selected when multiple PASes are present in the pre-mRNAs and how PAS selection can be regulated in a tissue- or developmental stage-specific manner. These questions are important as alternative polyadenylation (APA) is increasingly recognized as a critical mechanism for post-transcriptional gene regulation (13–18), and APA regulation impacts a variety of physiological processes including stem cell differentiation and cancer development (19–20). Although a number of protein factors have been shown to

*To whom correspondence should be addressed. Tel: +86 13002032215; Email: yaochguo@mail.sysu.edu.cn
Correspondence may also be addressed to Yongsheng Shi. Tel: +1 949 824 0358; Email: yongshes@uci.edu
†These authors contributed equally to this work as first authors.

regulate APA, it remains completely unknown whether any trans-acting RNAs may play a role in this process.

SnoRNAs are among the most abundant non-coding RNAs (ncRNAs) in the nucleus, and they play an important role in rRNA and snRNA modification, including uridine isomerization and ribose methylation (21–22). Based on the structural features, snoRNAs are classified into two families, C/D-box and H/ACA-box snoRNAs. SnoRNAs bind to target RNA sequences through base-pairing and recruit partner proteins for modification. It is important to point out that, for most snoRNAs, it has been difficult to experimentally identify their target sequences, at least in part due to the lack of complementarities of snoRNAs and their targets (23). Interestingly, there is accumulating evidence that snoRNAs could be assembled into non-canonical snoRNP particles and have a wide variety of functions, ranging from RNA silencing, pre-mRNA splicing to chromatin decondensation (24–28). Some snoRNAs, such as SNORD50A, a C/D-box snoRNA, has been previously implicated in a variety of cancers (29–33). Therefore, it is important to delineate snoRNA functions, both canonical and noncanonical.

In our attempt to systematically characterize the mammalian mRNA 3′ processing complex, we have unexpectedly found that a subset of snoRNAs physically associate with the pre-mRNA 3′ processing complex. To elucidate the potential function(s) of this association, we have focused on SNORD50A and showed that this U/A-rich snoRNA specifically binds to Fip1 and in turn blocks Fip1 interaction with PAS sequences *in vitro* and *in vivo*. Consistently, excess SNORD50A inhibits mRNA 3′ processing while its depletion leads to an increase in mRNA 3′ processing efficiency and higher expression of many mRNAs. Our data revealed a novel function for snoRNAs and provided the first evidence that trans-acting RNAs can regulate mRNA 3′ processing.

MATERIALS AND METHODS

Cell culture and transfection

HeLa cells were cultured in DMEM with 10% FBS. Modified ASOs (50 nM) targeting SNORD50A (ASO1: mU∗mU∗mU∗mU∗mU∗T∗T∗T∗T∗C∗A∗A∗C∗A∗G∗mA∗mA∗mG∗mU∗mU; ASO2: mC∗mU∗mC∗mA∗mG∗A∗A∗G∗C∗C∗A∗G∗A∗T∗C∗mC∗mG∗mU∗mA∗mA; first and last five nucleotides are 2′-*O*-methyl RNA, other 10 nucleotides are DNA; The backbone is linked by phosphorothioate bond) or siRNAs (50 nM) targeting Fip1 (CGAUGAAGAACGAUACAGATT) were transfected with Lipofectamine RNAimax (Life Technology), and the same with the corresponding nontarget controls. Lipofectamine 2000 (Life Technology) was used to transfect luciferase reporter plasmid pPASPORT and snoRNA overexpression plasmid pCMV-globin. Cells were harvested at the indicated time points for further analysis upon transfection.

RNA-biotin based pull down assay

RNA substrates were made by *in vitro* transcription using SP6 or T7 RNA polymerase and DNA templates (PCR product or linearized Plasmid; two annealed DNA oligos used for random RNA N75), and biotinylated at 3′ end using a biotinylation Kit (Thermofisher). HeLa NEs were made following the described protocol (34). Biotinylated RNAs were first bound to the streptavidin beads, and then incubated with HeLa NE in the polyadenylation conditions [see (7) for details] for 20 min, after biotin–streptavidine binding, washing, pull-down sample were heated (75° for 5 min) in 1× SSC buffer for elution, as suggested by the manual of the Dynabeads M280. The eluted samples were further subject to western blots and/or RNA purification for RNA-seq/RT-qPCR.

Gel shift assays

Gel shift assays were performed by incubating P32 labeled RNAs (detectable amount, ∼2 nM) with indicated amount of GST-Fip1-RRM (or together with cold RNAs in competition assay) in 15 μl of binding buffer [10 mM HEPES (pH 7.9), 50 mM NaCl, 0.5 mM $MgCl_2$, 0.1 mM EDTA, 5% glycerol, 1 mM ATP, 10 mM creatine phosphate, 5 mM β-mercaptoethanol, 0.25 mM PMSF, 0.7 μg of *Escherichia coli* tRNA, and 1.4 μg of BSA] at 30°C for 10 min. Samples were further loaded onto 5% nondenaturing PAGE gels. The radioactivity signals were analyzed by PhosphorImager (Typhoon FLA 7000).

Luciferase reporter assay

HeLa cells were harvested 24 h after transfection with pPASPORT plasmids. Luciferase activity was measured using Promega Dual-Luciferase Reporter kit and Berthold Sirius detection system.

In vitro mRNA 3′ processing assays

In vitro cleavage/polyadenylation and cleavage assays were performed using P32-radiolabeled SVL PAS RNA sequences and HeLa NEs following standard protocol as described elsewhere (8,11). Indicated amount of SNORA78 RNAs or SNORD50A was added in the assay in Figure 2A.

Northern blot and RT-qPCR

Total RNAs were prepared using Trizol reagent (Life Technology). Ten microgram of total RNAs or other indicated RNA samples were separated in 10% polyacrylamide–7 M urea gels and transfered onto positively charged nylon membrane, using a semi-dry transfer apparatus (Bio-Rad). Northern blot analysis of snoRNAs and U6 snRNA were performed using 5′-radiolabeled DNA probes (see Supplementary Table S5 for probe sequences) in ultrasensitive hybridization Buffer (ThermoFisher). Real-time PCR was performed in 96-well plates with LightCycler® 480 qPCR system (Roche). Briefly, cDNAs were produced from target RNAs with superscript III (Life Technology) and quantified by amplification using corresponding primers (see Supplementary Table S5), qPCR data were analyzed by ΔΔCt methods, and normalized by appropriate controls (si-NC/NC-ASO/Gapdh/28S rRNA gene).

Immnoprecipitation and western blots

All the antibodies for 3′ processing factors were from Bethyl Laboratories, other antibodies were from Abcam or Santa Cruz Biotechnology (Cat. number available on request). Nuclear extract was diluted with equal volume of Buffer D [20mM HEPES (pH 7.9); 100 mM NaCl; 1 mM MgCl2; 0.2 mM EDTA; 10mM β-ME; 0.5 mM PMSF], and supplemented with 0.1% NP-40. IPs were performed using Dynabeads with diluted NE and indicated antibodies. Immunoprecipitated samples were further subject for western blots or RNA extraction. For western blots, primary antibodies are the same as the ones for IP. For secondary antibodies, we used HRP-conjugated anti-mouse/rabbit (sigma) or TrueBlot anti-mouse/rabbit (eBioscience). ECL western blot system (ThermoFisher) was used to detect the signals.

RNA-seq

Strand-specific RNA-seq libraries were prepared following the described protocol (35) with the RNA-biotin based pull-down sample, and single-end squenceing (SE50) was carried out in Illumina Genome Analyzer II platform. Raw reads were trimmed by Trimmomatic (V0.32), and mapped to the human genome (hg19), allowing up to two mismatches using OSA (Omicsoft Sequence Aligner, version 7.0.1.20). Expression level estimation was reported as TPM (transcripts per million) value for each sample. Stringent cut-off parameters (average TPM $>$ 200, fold change $>$ 5) was applied to select highly reliable changes.

PAS-seq

Oligo(dT) priming based PAS-seq libraries were prepared following the 3PC protocol (36) with minor modification. Briefly, 15 μg total RNAs were fragmented, poly(A)+ transcripts were converted into cDNA with a primer that contains an oligo(dT) sequence, two primer binding sites for nested PCR and two BamHI restriction enzyme sites. cDNAs were size-fractionated, purified, and further subject to circularization and linearization. A limited number of PCR cycles were used to amplify the linearized cDNAs. Single-end sequencing (SE100) was carried out in Illumina Genome Analyzer II platform. Raw reads were trimmed, mapped to human genome (hg19), allowing up to two mismatches using Bowtie2 with the settings 'bowtie2 -p 28 -N 1 -k 1′. A poly(A) junction is where the 3′ end of the read maps. Internal priming issue was addressed as described previously (37). For the analysis shown in Figure 3, we first mapped and clustered the reads to Ensembl genes (GRCh37.75). Next we used Fisher's exact test to compare the ratio of the read counts of one PAS to the total read counts the same gene between two samples. PASes with an adjusted P-value $<$0.01 (adjusted by Benjamini–Hochberg method) were considered as differentially used. To create the scatter plot shown in Figure 3A, we selected two PASes with the smallest P-values for each gene with multiple PASes. The PAS closer to the transcript start site is designated as the proximal PAS and the other as distal. We then calculated the corresponding proximal/distal ratio. PAS pairs with an adjusted P-value $<$0.01 and $|\log_2$(control/SNORD50A KD-proximal/distal ratio)$| > 0.5$ are highlighted (8). For gene expression analysis in Figure 3B, all the normalized PAS-seq reads (RPM) for a gene were summed to stand for the expression level of this gene, and DESeq2 package was used to analyzed the gene expression change.

Fip1 iCLIP-seq

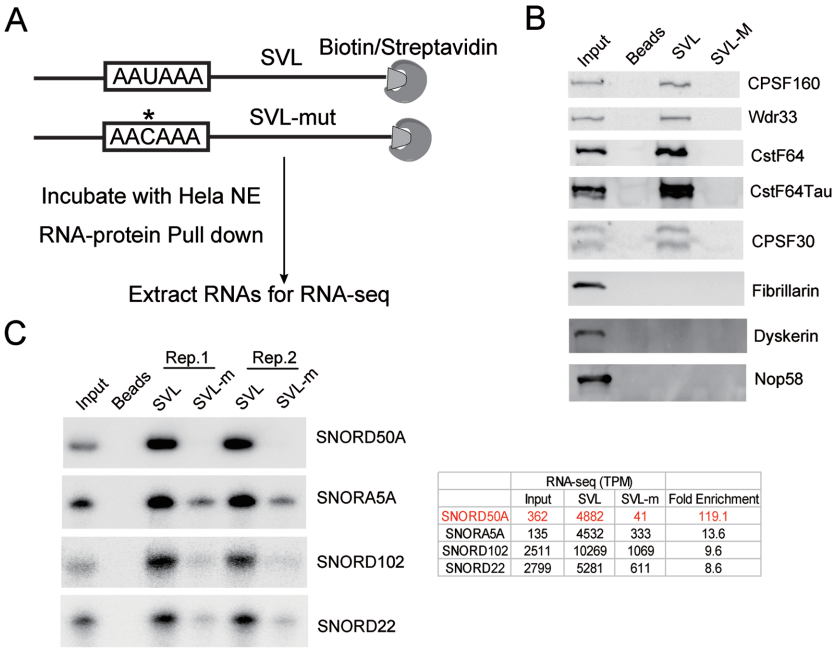
Fip1 iCLIP-seq was performed and analyzed essentially as previously described (20).

RESULTS

A subset of snoRNAs are associated with the mRNA 3′ processing complex

We have previously purified the human mRNA 3′ processing complex and comprehensively characterized its protein composition (7). Next we wanted to explore the possibility that there may be RNAs associated with this complex. To test this, we have purified the human mRNA 3′ processing complexes, extracted the associated RNAs, and subjected them to high-throughput sequencing analyses. Briefly, as illustrated in Figure 1A, we used SV40 late (SVL) PAS, a commonly used PAS RNA substrate for *in vitro* mRNA 3′ processing assay (7,8,11,20,38). A mutant RNA with a single point mutation (U to C) within the poly(A) signal AAUAAA was used as negative control as it does not allow the assembly of mRNA 3′ processing complex. The RNA substrates were first biotinylated at the 3′ end, and then bound to the streptavidin magnetic beads. After incubation with HeLa nuclear extract (NE) under cleavage/polyadenylation conditions, the RNA–protein complexes were purified by using biotin–streptavidin pull-down. Western blot analyses revealed that all the known 3′ processing factors, including Wdr33, CPSF30, CPSF160, CstF64, CstF64τ (8–10,12,39–41), were specifically associated with the wild type PAS RNA, but not with the mutant RNA (Figure 1B). These results are highly consistent with previous studies (7), and suggest that our biotin–streptavidin pull-down strategy is highly specific and suitable for identifying components of the mRNA 3′ processing complex.

Next, we extracted RNAs from the purified human mRNA 3′ processing complexes, prepared strand-specific RNA-Seq libraries and subjected them to deep-sequencing analyses using a previously described method (35). For each library, we obtained around 1 million reads uniquely mapped to the human genome. Using stringent cut-off parameters [average TPM $>$200, fold change (SVL/mutant) $>$ 5], we found 11 RNAs that are enriched in SVL RNA pull-down sample compared to the mutant RNA pull-down sample. (Supplementary Table S1). Remarkably, almost all the enriched RNAs (10 out of 11) corresponded to snoRNAs. Among them, nine are C/D-box snoRNAs and 1 is H/ACA-box snoRNA. Similarly, 73% of the enriched RNAs (64 out of 88) are snoRNAs if a less stringent cutoff value (Average TPM$>$100, Fold change $>$ 1.5) was applied. These results indicated that specific snoRNAs were associated with mRNA 3′ processing. Importantly, although we did not have replicates in the RNA-seq data, our follow-up



	RNA-seq (TPM)			
	Input	SVL	SVL-m	Fold Enrichment
SNORD50A	362	4882	41	119.1
SNORA5A	135	4532	333	13.6
SNORD102	2511	10269	1069	9.6
SNORD22	2799	5281	611	8.6

Figure 1. Identification of trans-acting RNAs associated with mRNA 3′ processing. (**A**) Schematic representation of the SVL RNA substrates used in the biotin–streptavidin pull-down assay. The AAUAAA hexamer in wild-type RNA substrate and AACAAA in mutant substrate (boxes) are shown. The asterisk is used to highlight the single nucleotide change. (**B**) Western blot analysis of known core 3′ processing factors and snoRNP proteins in the RNA-biotin based pull-down experiment. (**C**) The RNA-seq fold-change values for 4 snoRNAs that were enriched in wild-type SVL pull-down samples. Northern blots were performed to confirm the enrichment (Left panel).

northern blot (Figure 1C) or RT-qPCR analysis (Supplementary Figure S2A) has confirmed the fold changes for all of the nine randomly selected snoRNAs, suggesting that our sequencing analyses were highly reliable.

snoRNAs are known to assemble into RNPs with several specific protein partners (22). Given our surprising finding that specific snoRNAs were enriched in the pull-down sample, we next asked if the canonical snoRNP particles were captured under our purification conditions. To this end, we have performed Western blotting analyses of our purified mRNA 3′ processing complexes for known core snoRNP proteins such as dyskerin, NOP58 and fibrillarin (22,42,43). Consistent with previous proteomic analyses results (7), we failed to detect these proteins in the pull-down samples (Figure 1B), indicating that the snoRNAs associated with mRNA 3′ processing complexes were not assembled into the canonical snoRNPs.

SNORD50A inhibits SVL PAS 3′ processing in vitro and in vivo

We next asked whether mRNA 3′ processing complex-associated snoRNAs play a role in mRNA 3′ processing. To this end, we have focused on SNORD50A, which is the most enriched snoRNA in the SVL mRNA 3′ processing complex (Figure 1C). Interestingly, SNORD50A has been reported to be associated with cancers (29–33).

We first tested the effect of SNORD50A on mRNA 3′ processing *in vitro*. To this end, we carried out *in vitro* cleavage/polyadenylation assay in the presence of increasing amount of SNORD50A. SNORA78 was used as a negative control as it does not associate with mRNA 3′ processing complex (Supplementary Table S1). As shown in Figure 2A and Supplementary Figure S1A, SNORD50A significantly inhibited SVL cleavage/polyadenylation efficiency in a dose-dependent manner while SNORA78 had no effect, suggesting that SNORD50A can specifically suppress mRNA 3′ processing at SVL PAS *in vitro*.

To test the role of SNORD50A *in vivo*, we over-expressed SNORD50A and several control snoRNAs (SNORA16A, SNORD20 and SNORA78) that are not associated with mRNA 3′ processing complex in HeLa cells using a previously reported vector system (44,45; Supplementary Figure S1B). Briefly, in this expression system, snoRNA insert was processed from the second intron of human β-globin gene, which had been placed downstream of the cytomegalovirus (CMV) promoter. Our RT-qPCR analyses showed that transfection of pCMV-globin-snoRNA expression construct led to a ∼50-fold increase of corresponding snoRNA compared to the transfection of a control vector (Supplementary Figure S1B). Additionally, we have also validated the overexpression of SNORD50A through Northern blot analysis (Supplementary Figure S1B). To assess the effect of snoRNA overexpression on mRNA 3′ processing efficiency *in vivo*, we took advantage of a previously described bicistronic luciferase reporter construct, pPAS-PORT ((8,9,20); Supplementary Figure S1C). Briefly, in this assay system, efficient mRNA 3′ processing at the tested PAS leads to high expression of Renilla luciferase gene and low expression of Firefly luciferase gene, while poor mRNA 3′ processing results in the opposite mode of gene expression. Therefore the Renilla/Firefly ratio provides a quantitatively measurement of the cleavage/polyadenylation efficiency at the tested PAS. Our results showed that overex-

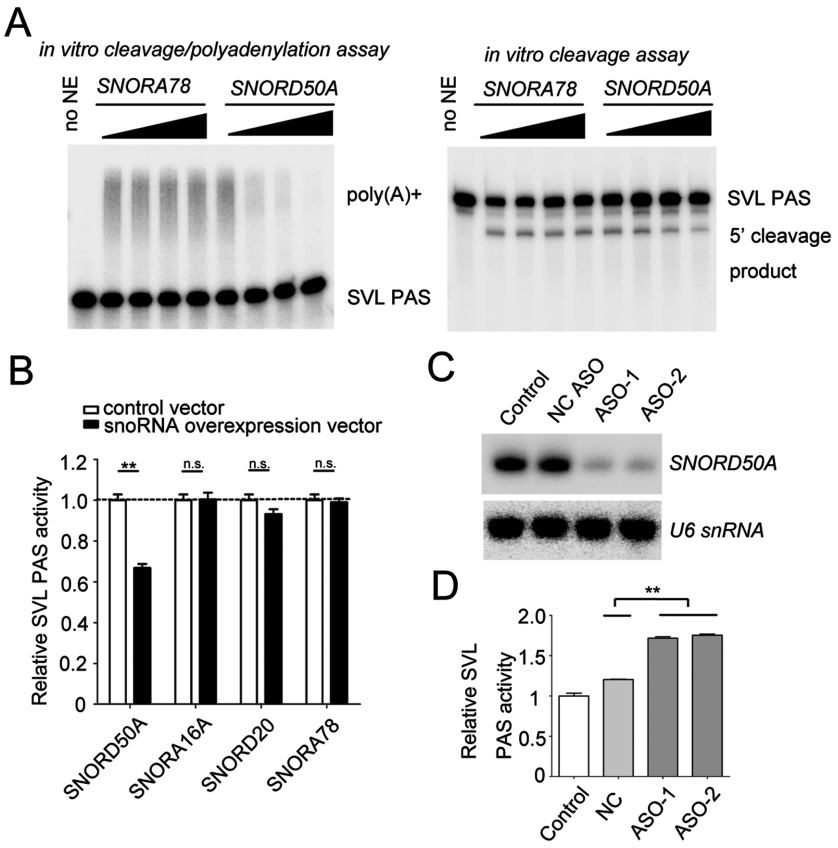


Figure 2. Functional characterization of SNORD50A in SVL PAS 3′ processing. **(A)** *In vitro* cleavage /polyadenylation (left) and cleavage (right) assays using SVL PAS RNA substrate and nuclear extract from HeLa cells in the presence of SNORA78 RNA oligo (0, 5,10, 15 μM) or SNORD50A (0, 5, 10, 15 μM). pre-mRNA, polyadenylated mRNA, 5′ cleavage product are indicated. (**B**) Luciferase reporter assays to determine the impact of individual snoRNA overexpression on SVL PAS 3′ processing. pCMV-globin control vector and pCMV-globin-snoRNA expression construct were transfected into HeLa cells, after 2 days, SVL PAS activities were determined by transfecting pPASPORT-SVL plasmid construct into HeLa cells and measuring the *Rluc*/*Fluc* ratio 1 day after the transfection. The ratios were normalized by the value of SVL PAS activity in HeLa cells transfected with pCMV-globin control vector, and plotted as mean±S.E.M. (n = 3). Student's t-test was used to determine the significance of the change. n.s.: not significant; $^{**}P < 0.01$. (**C**) Northern blot analysis of SNORD50A expression in control HeLa cells and HeLa cells transfected with negative control (NC) ASO, 2 ASOs (ASO-1, ASO-2) targeting SNORD50A. Total RNAs were harvested 2 days after transfection. U6 snRNA expression serves as loading control. (**D**) Luciferase reporter assays to determine the impact of SNORD50A depletion on SVL PAS 3′ processing. Control and SNORD50A ASOs were transfected into HeLa cells, after 2 days, SVL PAS activities were determined by transfecting pPASPORT-SVL plasmid construct into HeLa cells and measuring the *Rluc*/*Fluc* ratio 1 day after the transfection. The ratios were normalized by the value of SVL PAS activity in control HeLa cells and plotted as mean ± S.E.M. ($n = 3$). Student's t-test was used to determine the significance of the change. $^{**}P < 0.01$.

pression of SNORD50A led to significant inhibition of SVL 3′ processing efficiency compared to that of other control snoRNAs (Figure 2B), suggesting that SNORD50A can specifically suppress mRNA 3′ processing at SVL PAS *in vivo*.

To further study the function of SNORD50A *in vivo*, we next depleted SNORD50A in HeLa cells using antisense oligonucleotides (ASO), which has been well established as a specific and efficient tool for depleting nuclear ncRNAs, including snoRNAs (46,47). Consistent with previous results, SNORD50A was specifically depleted by 90% upon ASO transfection (46) (Figure 2C). Using this system, we first assessed the effect of SNORD50A depletion on mRNA 3′ processing efficiency using the aforementioned reporter. Our results showed that the mRNA 3′ processing efficiency for the SVL PAS was moderately increased upon SNORD50A knockdown (Figure 2D). snoRNAs are involved in rRNA modification and snoRNA depletion could affect translation. To rule out any indirect effect in our reporter assay, we also quantified Renilla and Firefly mRNA levels in our reporter assays by RT-qPCR. Consistently, the ratio of Renilla/Firefly mRNA is significantly higher in SNORD50A knockdown sample (Supplementary Figure S1C). Therefore, our *in vivo* and *in vitro* analyses provided consistent and strong evidence that SNORD50A negatively regulates mRNA 3′ processing at the SVL PAS.

SNORD50A modulates APA and expression level of endogenous mRNAs

A number of core mRNA 3′ processing factors or their associated proteins have been shown to regulate APA and/or transcript abundance *in vivo* and their impacts are gene-specific (8,9,11,20,48,49). Next, we asked whether SNORD50A can impact global APA and gene expression profiles. To this end, 3PC (3′ PolyA site mapping using cDNA circularization), a previously characterized PAS-

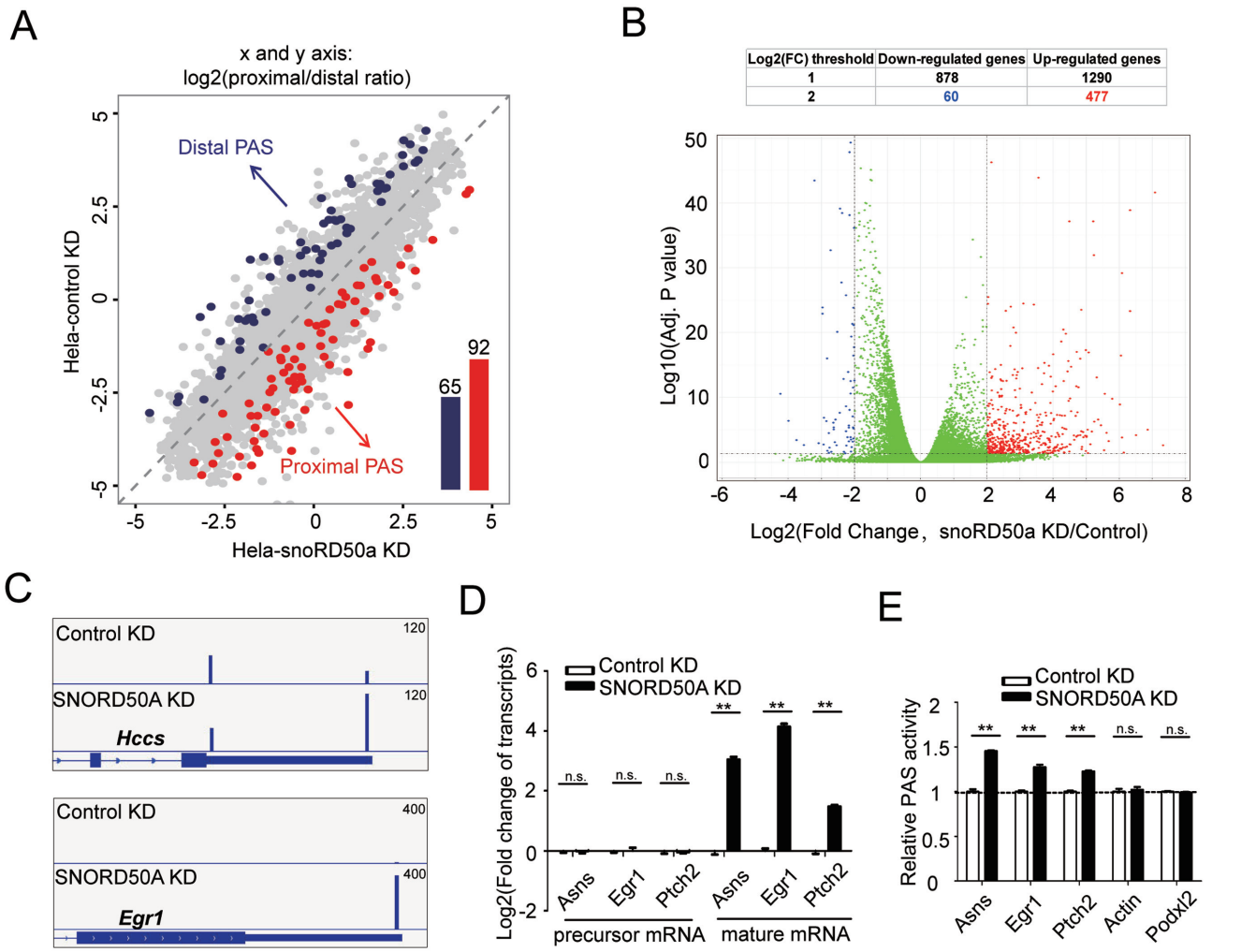


Figure 3. Effect of SNORD50A depletion on the 3′ processing and expression level of endogenous poly(A)+ transcripts. (**A**) PAS-seq analysis of APA in HeLa cells 2 days after they were transfected with NC ASO or SNORD50A ASO (ASO-1 and ASO-2 as a group), Log2(proximal/distal ratio) are plotted for HeLa cells transfected with NC ASO (y-axis) and SNORD50A KD ASO (x-axis). Statistically significant changes are highlighted in blue (proximal to distal shift) and red (distal to proximal shift). The numbers of APA changes are shown in the column graph. (**B**) Volcano plot showing the expression level change of poly(A)+ mRNA in HeLa cells 2 days after they were transfected with NC ASO or SNORD50A ASO (ASO-1 and ASO-2 as a group). Highly reliable changes ($P<0.05$, fold change>4) were colored with red (up-regulated in SNORD50A KD cells in comparison to NC KD cells) or blue (down-regulated in SNORD50A KD sample in comparison to NC KD cells), green dots shows the changes either not statistically significant (P>0.05) or less reliable (fold change<4). (**C**) IGV tracks showing the PAS-seq results for *Hccs, Egr1* genes in HeLa cells with NC (top) or SNORD50A KD (bottom). (**D**) RT-qPCR analysis monitoring the relative expression levels of 3 mature and precursor transcripts in HeLa cells 2 days after they were transfected with NC ASO or SNORD50A ASO. Fold changes are relative to *Gapdh* and normalized to NC ASO control, plotted as mean±S.E.M. (n = 3). Student's t-test was used to determine the significance of the change. n.s.: not significant; ** $P<0.01$. (**E**) Luciferase reporter assays are similar to Figure 2D except that the targets are 5 endogenous mRNAs PASes instead of SVL PAS. Student's t-test was used to determine the significance of the PAS activity change, ** $P<0.01$.

seq (PolyA Site Sequencing) method capable of quantifying polyA+ transcripts and gene expression at transcriptomic level (36), was carried out in negative control (NC) and SNORD50A knockdown HeLa cells (see Methods for details). Our PAS-seq analyses detected significant APA changes in 157 genes between control and SNORD50A-depleted cells. Among them, 65 genes showed a shift toward distal PASes while 92 genes showed a shift toward proximal PASes (Figure 3A; Supplementary Table S3). In addition to APA changes, the mRNA levels of 1290 genes were up-regulated and 878 genes down-regulated upon SNORD50A depletion by more than 2-fold, including 477 genes up-regulated and 60 genes down-regulated by >4-fold (Figure 3B; Supplementary Table S4). Figure 3C shows representative examples of the APA (top panel) and gene expression changes (bottom panel). Gene ontology analysis of the highly up-regulated genes [$\log_2$(fold change) > 2; RPM > 100] revealed a significant enrichment of genes that function in cell stress/proliferation/apoptosis (Supplementary Figure S3A), while the same analysis returned no enriched functional groups in down-regulated genes. These results suggest that SNORD50A regulates the expression of a subset of genes and its effect is largely negative.

mRNA 3′ processing complex-associated snoRNAs directly bind to Fip1

To begin to characterize the mechanisms for snoRNA-mediated regulation of mRNA 3′ processing, we set out to identify the direct binding partner(s) of these snoRNAs within the mRNA 3′ processing complex. To this end, we immunoprecipitated (IP) many known 3′ processing factors, extracted their associated RNAs, and performed RT-qPCR analysis using primers specific for the four most highly enriched snoRNAs found in the mRNA 3′ processing complex. Interestingly, we found that all four snoRNAs were most enriched in Fip1 IP sample (Supplementary Figure S2B). In addition to Fip1, which is a subunit of the CPSF complex, snoRNAs were also associated with some other CPSF subunits with moderate efficiency (Supplementary Figure S2B). To further test the association between CPSF and snoRNAs, we have performed pull-down assays with biotinylated-snoRNAs and HeLa NE. Our results showed that CPSF subunits were greatly enriched in biotinylated-snoRNAs pull-down sample compared to the biotinylated random RNAs and SNORA78, a snoRNA that is not associated with mRNA 3′ processing complex based on our analyses (Figure 4A; Supplementary Table S1). These data suggest that a subset of snoRNAs binds to mRNA 3′ processing complex at least in part through association with the CPSF complex.

To determine which CPSF subunit(s) directly bind to these four snoRNAs *in vivo*, we performed CLIP (UV cross-linking and immunoprecipitation) using antibodies against five CPSF subunits, extracted RNAs, and carried out RT-qPCR to detect specific target RNAs. It has been well established that CLIP detects direct protein–RNA interactions *in vivo* (50). Additionally, to further ensure that we were detecting specific protein–RNA interactions and not RNA interactions of other associated proteins, we denatured the cell lysates by heating it at 75°C for 5 min in the presence of SDS before immunoprecipitating protein–RNA complexes (Supplementary Figure S2C). Under such denaturing conditions, all protein–protein interactions were disrupted as evidenced by the absence of other CPSF subunits in Fip1 IP samples (Supplementary Figure S2C). Under this condition, we found that four snoRNAs were significantly enriched in the Fip1 IP sample compared to the IP samples of other CPSF subunits (Figure 4B). Consistently, examination of previous published CLIP-seq (UV cross-linking and immunoprecipitation followed by high-throughput sequencing) data of 3′ processing factors in HEK293 cells showed that CPSF subunits bind to all four snoRNAs and that Fip1 has the highest CLIP-seq signals for SNORD50A and snoRD22 (51) (Supplementary Table S2). In keeping with our earlier results, we did not detect any of the canonical snoRNP proteins in Fip1 IP sample (Supplementary Figure S2D). Taken together, these results strongly suggest that a subset of snoRNAs associate with mRNA 3′ processing, at least in part, by directly binding to Fip1.

SNORD50A blocks Fip1–PAS interaction

We next wished to test whether and how SNORD50A–Fip1 interaction impacts mRNA 3′ processing activities. We first asked whether SNORD50A modulates Fip1 interactions with RNAs or other proteins within the mRNA 3′ processing complex. To test this, we incubated SVL pre-mRNA substrate with control and SNORD50A-depleted HeLa NE under cleavage/polyadenylation conditions, immunoprecipitated Fip1, and extracted RNAs or proteins from IP samples. Our western blot analyses of the IP samples detected similar amounts of CPSF160, CPSF73 and CPSF30, suggesting SNORD50A depletion had little effect on the interactions between Fip1 and other 3′ processing factors (Figure 4C, top panel, western blot). In contrast, as shown in Figure 4C (bottom panel), we found that more SVL RNA substrates were recovered by Fip1 IP in SNORD50A depleted samples, suggesting that SNORD50A inhibits Fip1 interaction with the SVL PAS RNA.

How does SNORD50A inhibit the interaction between Fip1 and SVL pre-mRNA? We envisioned two scenarios: first, SNORD50A binds to SVL PAS through base-pairing and the resultant RNA duplex inhibits Fip1 binding to SVL. Second, SNORD50A may interact with Fip1, thereby blocking Fip1 interaction with SVL PAS. To test the first hypothesis, radiolabeled SVL PAS RNA sequence were biotinylated at 3′ end, and incubated with radiolabeled SNORD50A under cleavage/polyadenylation conditions. biotin–streptavidin pull-down assay showed that SVL PAS did not interact with SNORD50A *in vitro* (Figure 4D, lane 3). In contrast, the anti-sense RNA (of the same length as SNORD50A) to a portion of SVL PAS was able to bind to SVL PAS, while the sense sequence of the same region serves as a negative control (Figure 4D, lanes 1 and 2). These data suggest that SNORD50A does not hybridize to SVL PAS RNA.

We next tested the second hypothesis. Previous reports have suggested that Fip1 binds to U-rich sequence upstream of AAUAAA (20,38,51). A sequence comparison showed that the SNORD50A sequence is similar to a portion of the SVL PAS sequence upstream of the AAUAAA hexamer (Figure 3B), raising the possibility that SNORD50A might compete with SVL PAS RNA for binding to Fip1. Previous studies have shown that Fip1 binds to RNA via its arginine-rich C terminal region (Fip1-C) (38). To test this idea, we performed gel shift assay using recombinant Fip1-C (GST-Fip1-C) and SVL PAS RNA. Indeed, GST-Fip1-C binds to both SNORD50A and SVL PAS upstream element (USE). We next performed a competition assays by incubating GST-Fip1-C, radio-labeled SVL PAS RNA, and increasing amounts of unlabeled SNORD50A RNA as competitors. Our results showed that SNORD50A, but not SNORA78, could outcompete SVL USE for binding to Fip1-C (Figure 4E, left and right picture).

Our above results suggest that SNORD50A directly binds to Fip1 and serves as a competitive inhibitor for SVL PAS 3′ processing. Intriguingly, SNORD50A was identified within purified mRNA 3′ processing complex. How does this inhibitory snoRNA incorporated into the mRNA 3′ processing? Considering our Fip1 IP and RT-qPCR data, we hypothesize that SNORD50A binds to Fip1, which is assembled into mRNA 3′ processing complex through its interactions with other CPSF subunits. To test this, we performed the aforementioned RNA-biotin pull-down experiments using SVL PAS and nuclear extracts from HeLa cells

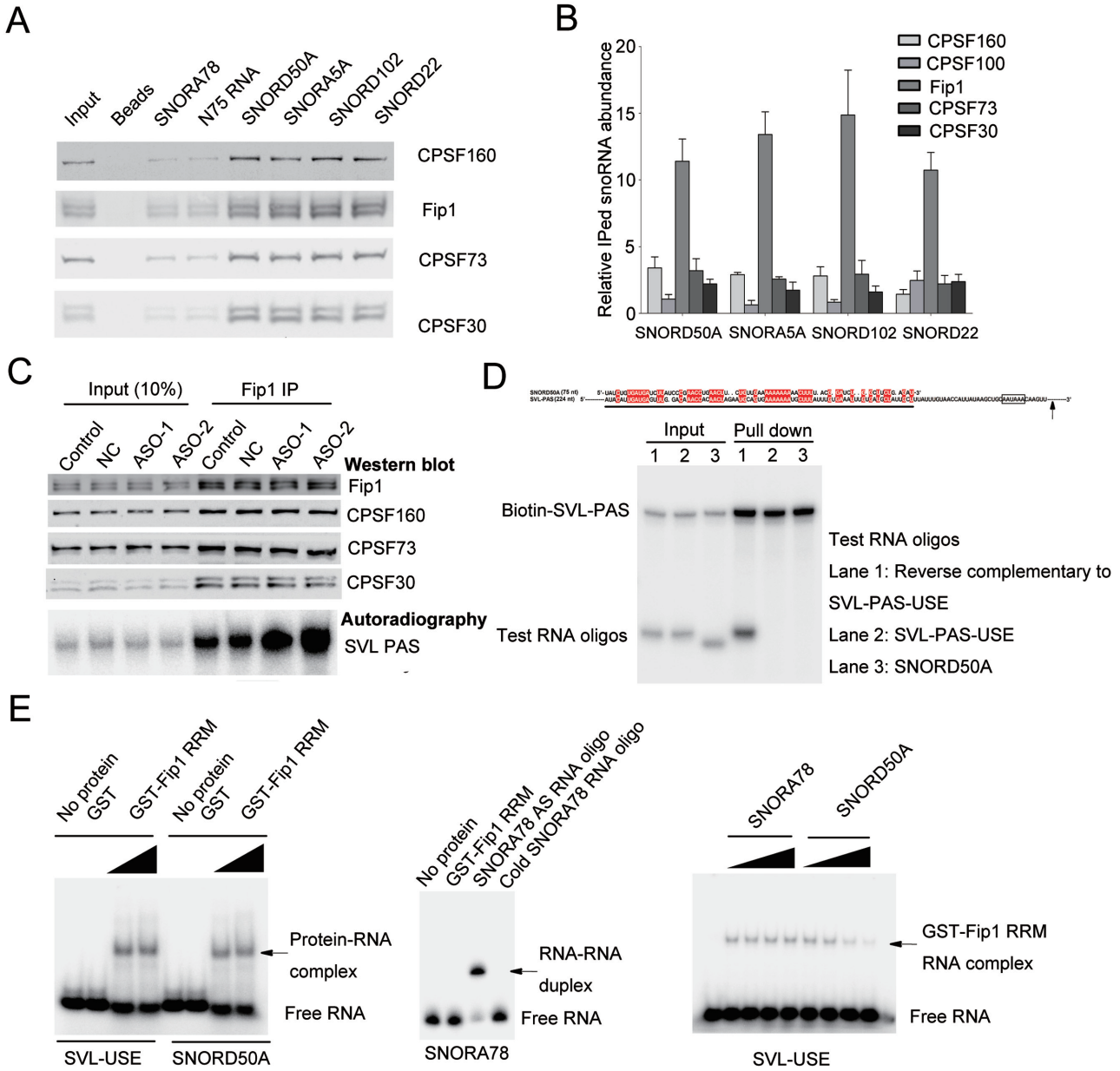


Figure 4. Mechanism study of SNORD50A-mediated SVL PAS 3′ processing regulation. (**A**) snoRNAs interact with CPSF in an *in vitro* binding assay. Control RNAs and four snoRNAs were biotinylated at 3′ end and incubated with HeLa NE in polyadenylation conditions, after binding, elution, biotin–streptavidin pull-down sample were subject to western blot analysis using antibodies against indicated CPSF subunits. (**B**) RT-qPCR analysis of the relative abundance of four snoRNAs in immunoprecipitated samples, IP was performed using 50 μl HeLa nuclear extract and 2 μg indicated antibodies. 10% input HeLa NE was used as normalization control. The IP step was performed following CLIP method (50) with modification (see result section for details). (**C**) Autoradiography of the P32 labeled SVL pre-mRNA (bottom) and Western blot analysis of CPSF subunits (top) from immunoprecipitation using Fip1 antibodies with control HeLa cell nuclear extracts (control) or nuclear extracts from NC ASO treated HeLa cells (NC), or nuclear extracts from SNORD50A ASO treated HeLa cells (ASO1, ASO2). P32-labeled SVL pre-mRNA were incubated with nuclear extracts under polyadenylation conditions for 20 minutes before immunoprecipitation. (**D**) RNA–biotin based pull-down assay showing SNORD50A did not bind to SVL RNA under polyadenylation conditions. Pairwise sequence alignment between SNORD50A and SVL PAS RNA sequences using CLUSTALW program (top). Same nucleotides were colored with red. AAUAAA hexamer and cleavage site within SVL PAS are shown in box and arrow respectively. SVL USE sequence marked with black line was used for gel shift, competition binding assay and pull-down assay. P32 radiolabeled SVL PAS was biotinylated at 3′ end, and incubated with same P32 radiolabeled SNORD50A or control RNA sequences (SVL USE serves as negative control, while its reverse complementary sequence serves as positive control) under buffer conditions compatible for polyadenylation. Pull-down samples were washed, eluted, and resolved by denaturing PAGE. (**E**) Gel mobility shift assays using GST-Fip1-RRM (25, 50 μM) and the P32 labeled RNA sequences (left), GST (50 μM) or no protein was used as control. For the RNA competition binding assay (right), SNORA78 RNA oligos (0, 5, 10, 15 μM) or SNORD50A (0, 5, 10, 15 μM) was present in the gel mobility shift assay using GST-Fip1-RRM (10 μM) and P32-labeled SVL USE RNA sequence. As negative control (middle), SNORA78 were not able to form stable complex with GST-Fip1-RRM.

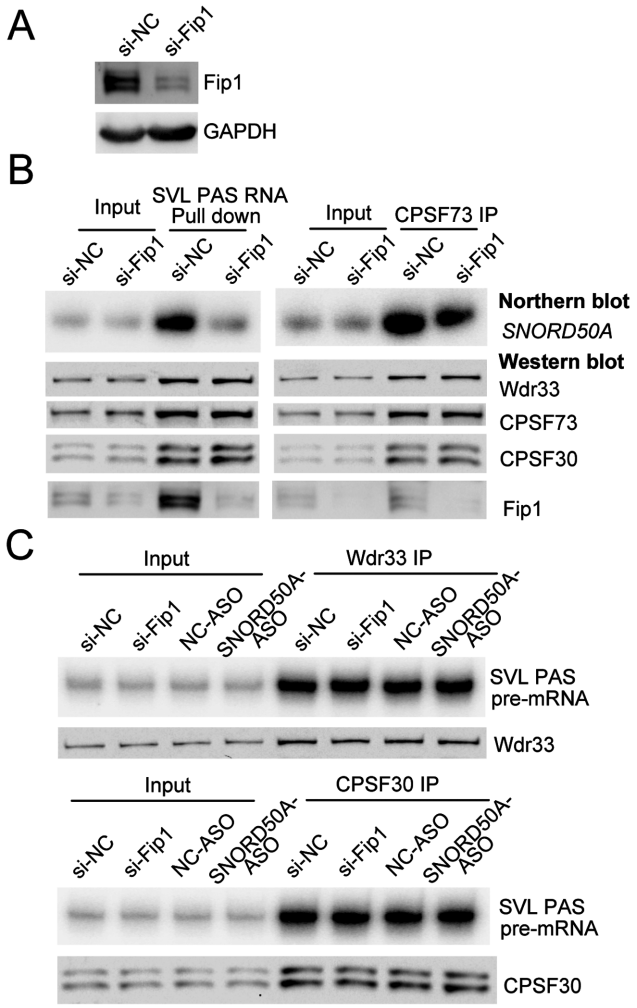


Figure 5. Mechanism study of SNORD50A assembly into CPSF and its effect on the interaction of CPSF with SVL PAS. (**A**) Western blot analysis of Fip1 expression in HeLa cells treated with negative control siRNA and siRNA targeting Fip1. (**B**) Northern blot analysis of SNORD50A (top) and western blot analysis of CPSF subunits (bottom) in SVL RNA–biotin based pull-down samples (left panel) and CPSF73 immunoprecipitated samples (right panel). Nuclear extracts from HeLa cells treated with NC siRNA or Fip1 siRNA were used in the pull-down assays. (**C**) Autoradiography of the P32 labeled SVL pre-mRNA (top) and western blot analysis of WDR33/CPSF30 from immunoprecipitation using WDR33/CPSF30 antibodies with nuclear extracts from NC ASO/siRNA treated HeLa cells, or nuclear extracts from SNORD50A ASO/Fip1 siRNA treated HeLa cells. P32 labeled SVL pre-mRNA were incubated with nuclear extracts under polyadenylation conditions for 20 min before immunoprecipitation.

transfected with control siRNA or Fip1 siRNA (20) (Figure 5A). Northern blot analysis showed that a lower amount of SNORD50A was assembled on SVL PAS when nuclear extracts from HeLa cells treated with Fip1 siRNA was used (Figure 5B, left picture, top). Consistently, CPSF73 IP experiments using HeLa NEs showed that less amount of SNORD50A was present within CPSF upon Fip1 depletion (Figure 5B, right picture, top). Interestingly, Fip1 KD did not affect the integrity of the CPSF complex or the recruitment of CPSF to SVL PAS (Figure 5B, left and right pictures, bottom). These results suggest that Fip1 indeed mediates the recruitment of SNORD50A to CPSF and in turn the mRNA 3′ processing complex.

We next investigated whether SNORD50A also blocks the interaction of other CPSF subunits with SVL PAS. To address this, we performed similar experiment as described in Figure 4C. Briefly, radiolabeled SVL PAS RNA was incubated with HeLa NE and RNA-IP was carried out using antibodies against WDR33 or CPSF30, we focused on these two proteins as recent studies demonstrated that they directly bind to AAUAAA within mammalian PAS (10,12). Our results showed that knockdown of either SNORD50A or Fip1 had no significant effect on the association between WDR33 and CPSF30 with SVL PAS (Figure 5C). Together, these data suggest that SNORD50A specifically blocks Fip1 interaction with SVL PAS.

SNORD50A regulates 3′ processing of endogenous mRNAs at least in part via modulation of the interaction of Fip1 with PASes

To determine whether and how snoRNAs affect endogenous mRNA 3′ processing, we next wished to select several direct target sites for SNORD50A from the 477 up-regulated genes in SNORD50A KD cells and characterize the mechanism for SNORD50A-mediated regulation of cellular PASes. Previous studies have indicated that, for genes containing more than one PAS, the PAS selection could be affected by a variety of factors, including the relative intrinsic strength of the different PASes, the length between the alternative PASes, and splicing (14,18,20,52). Therefore, to minimize the number of factors that may complicate our mechanistic study, we focused on 106 genes that only have a single PAS (Supplementary Table S4).

Next we selected several mRNA targets among the top candidates on the list for further verification. Consistent with PAS-seq data, elevated expression of the three selected mRNAs upon SNORD50A depletion were confirmed by RT-qPCR analysis (Figure 3D, right panel). The increased mRNA levels of these genes could be due to lower mRNA turnover rate and/or higher transcriptional activity. To distinguish between these possibilities, we first employed actinomycin D chase technique to measure the half-lives of the three mRNAs after SNORD50A depletion. No significant differences were observed between control and SNORD50A knockdown cells (Supplementary Figure S3B), indicating that SNORD50A did not affect the turnover rates of these mRNAs. To check the effect of snoRNAs on transcriptional activity, we monitored the nascent transcript levels by measuring the levels of unspliced transcripts from these genes. Our results showed they were not significantly altered (Figure 3D, left panel), indicating that the synthesis of nascent transcripts was not affected. Therefore, we hypothesized that the up-regulation of the three mRNAs could be due to more efficient mRNA 3′ processing after SNORD50A depletion. To test this idea, we measured the mRNA 3′ processing efficiency at the PASes of the snoRNA target genes using the bicistronic luciferase reporter as described in Figure 2B. Indeed, higher mRNA 3′ processing activity was observed for all three PASes in SNORD50A knockdown cells (Figure 3E). In contrast, PAS activity of *beta-actin* and *podxl2* genes were not af-

fected by SNORD50A depletion. These data strongly suggest that SNORD50A inhibits the 3′ processing of specific PASes *in vivo*.

We next wished to further characterize the effect of SNORD50A on these endogenous PASes. First, we synthesized these PASes (–100 to +100 nt relative to the cleavage site) *in vitro* and performed biotin–RNA IP using HeLa NE as described earlier. As expected, these biotinylated-RNAs efficiently pulled down CPSF factors and SNORD50A (Figure 6A). These results indicated that SNORD50A was physically associated with the mRNA 3′ processing complexes assembled on these endogenous PAS sequences. Second, we performed pair-wise sequence alignment between SNORD50A and these PAS regions. Interestingly, SNORD50A is highly U/A-rich and shows high sequence similarity to a region near the poly(A) signal 'A(U/A)UAAA' (Supplementary Figure S3C). Third, SNORD50A and the homologous regions within these PASes could compete for binding to Fip1, as demonstrated by the competitive gel mobility shift assay (Supplementary Figure S3C). Finally, at the genome-wide level, SNORD50A shows significantly higher sequence homology with PAS sequences of up-regulated genes than with control mRNA sequences (Supplementary Figure S3D and E).

The above results, together with the analysis on SVL PAS RNA substrate, prompted us to investigate whether SNORD50A systematically affects Fip1–mRNA interaction at PAS region in vivo. To address this, we performed Fip1 iCLIP-seq in control and SNORD50A KD HeLa cells. Consistent with previous reports (20,38), our results showed that Fip1 preferentially binds to RNA sequences upstream of A(A/U)UAAA hexamer and cleavage/polyadenylation sites in control HeLa cells. In SNORD50A KD cells, however, Fip1 RNA binding sites were more broadly distributed and the median position was shifted 15 nucleotides relative to A(A/U)UAAA (Figure 6B; Supplementary Figure S4A). Moreover, significantly higher Fip1 iCLIP-seq signals were observed surrounding PAS sequences upon SNORD50A KD (Figure 6C), as seen in *Ptch2* mRNA (Figure 6D). These results are consistent with our *in vitro* analysis on SVL substrate and supports a model wherein SNORD50A blocks the interaction of Fip1 with PASes and suppresses mRNA 3′ processing and the expression level of target mRNAs.

To understand how higher Fip1-PAS interaction contributes to higher gene expression levels for a subset of genes, we first divided expressed genes into three groups based on their fold change values between SNORD50A knockdown and control cells, namely up-regulated, unchanged, and down-regulated genes. Interestingly, after SNORD50A knockdown, significant higher Fip1 iCLIP-seq signals were detected surrounding PAS regions for all three groups (Supplementary Figure S4B). In search for sequence features that may explain this specific mode of gene regulation, we examined the PAS sequences of the 100 most up-regulated genes. Our results showed these PASes are more U-rich around the cleavage/polyadenylation site (–100 to 100 nt) in comparison to the PASes of other groups, with the most significant U-rich observed at region (–100 to –30 nt) (Figure 6E; Supplementary Figure S4C). As Fip1 is known to predominantly interacts with U rich upstream element of PASes (20,38), this result is consistent with our model that SNORD50A regulates mRNA 3′ processing and expression of mRNAs through modulating Fip1–RNA interactions.

DISCUSSION

In this study, we have provided, to the best of our knowledge, the first experimental evidence that snoRNAs are involved in regulating mRNA 3′ processing. Through detailed functional studies of a specific snoRNA, SNORD50A, we showed that the U/A-rich SNORD50A could compete with specific PAS *cis*-element for binding to Fip1, thus serving as a negative regulator in target PAS 3′ processing. As such SNORD50A controls APA and expression level of target mRNAs *in vivo*. A schematic model of SNORD50A-mediated regulation of mRNA 3′ processing is presented in Figure 7. These results not only provided new insights into the mechanism of mRNA 3′ processing, but also extended our understanding of the functions of snoRNAs.

Several early studies provided evidence that the mRNA 3′ processing complex might contain an snRNA (53,54). For example, micrococcal nuclease treatment of nuclear extract abolishes mRNA 3′ processing activity, indicating that one or more RNAs are required for mRNA 3′ processing. A later study, however, showed that adding purified *E. coli* RNAs restored mRNA 3′ processing to micrococcal nuclease-treated nuclear extract (55). These studies indicate that some RNA component may play a role in mRNA 3′ processing, but no specific RNA seems to be required for mRNA 3′ processing. In this report, we showed that several snoRNAs were associated with Fip1 in the mRNA 3′ processing complex. Our functional characterization of SNORD50A, including the analysis from snoRNA overexpression and luciferase reporter assays (Figure 2), suggest that CPSF-interacting snoRNAs predominantly serve as negative regulators in a sequence-dependent manner. Therefore, we conclude that snoRNAs are not essential for mRNA 3′ processing, but play a regulatory role.

It has been increasingly recognized that both canonical and non-canonical snoRNPs exist in cells (22,28,56–60), and thus the same snoRNAs could be assembled into different protein complexes for distinct functions. Indeed, accumulating evidence suggests that snoRNAs have a wide variety of noncanonical functions, ranging from RNA silencing, pre-mRNA splicing to chromatin decondensation (24–28). However, the dual functionality of snoRNAs may complicate functional studies. For example, we showed that SNORD50A depletion using ASO caused APA and mRNA level changes in many target genes. Given the function of snoRNAs in rRNA modification, it is conceivable that some observed effect in SNORD50A-depleted cells could be due to ribosome-associated translation defects. However, our data strongly argue against this possibility. First, SNORD50A is physically associated with specific mRNA 3′ processing factors, and is assembled into mRNA 3′ processing complexes of specific PASes (Figures 4B and 6A). Second, the association between snoRNAs and mRNA 3′ processing complexes is highly specific. For example, our results showed that SNORD50A is associated with the mRNA 3′ processing complex assembled on the SVL

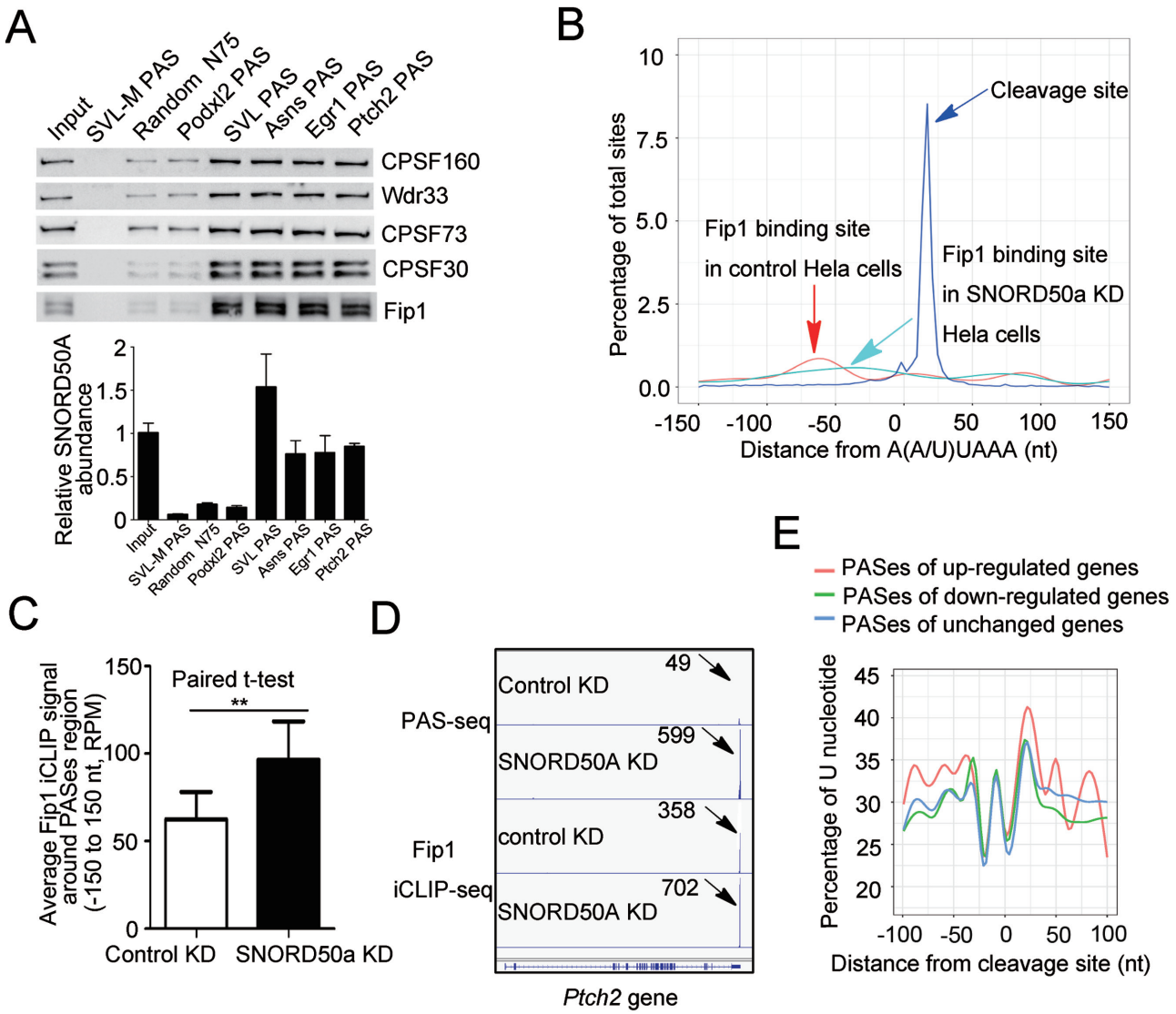


Figure 6. Mechanism study of SNORD50A-mediated endogenous mRNA 3′ processing regulation. (**A**) Endogenous mRNA PAS sequences interact with CPSF in an *in vitro* binding assay. Control RNAs (SVL-M PAS: negative control; SVL PAS: positive control; random N75: random technique control, Podxl2 PAS: non-target PAS control) and three mRNA PAS RNA sequences were biotinylated at 3′ end and incubated with HeLa NE in polyadenylation conditions, after binding, elution, biotin–streptavidin pull-down sample were subject to western blot analysis using antibodies against indicated CPSF subunits, and RNA purification for SNORD50A semi-quantification by RT-qPCR. (**B**) Distribution of Fip1 binding sites (red line: based on Fip1 iCLIP-seq data in control KD HeLa cells; green line: based on Fip1 iCLIP-seq data in SNORD50A KD HeLa cells) and cleavage sites (based on PAS-seq data) relative to closest upstream A(A/U)UAAA element. Position 0 stands for the 3′ end of A(A/U)UAAA. (**C**) Paired Student's *t*-test was used to compare the relative Fip1 binding frequency across PAS region (–150 nt to 150 nt region relative to cleavage site) for all the 3106 mRNAs (Fip1 iCLIP+ at PAS region) in HeLa cells upon SNORD50A KD. y-axis shows the number of total Fip1 iCLIP reads (normalized by RPM) across the PAS region. Bar graphs are plotted as mean ± SEM ($n = 3106$). $^{**}P < 0.01$. (**D**) PAS-seq and Fip1 iCLIP-seq tracks on the *Ptch2* gene. Normalized RPM values are shown for each sample (NC and SNORD50A KD HeLa cells) in the track. (**E**) Genes were divided into three groups based on $\log_2$(FC) value, namely up-regulated genes, down-regulated genes, unchanged genes (shown in red, green and blue, respectively). Distribution of percentages of U nucleotide sequence across PAS site region (–100 nt to 100 nt relative to cleavage/polyadenylation site) for the 3 groups of genes were plotted.

PAS. Interestingly, however, SNORD50B, a homolog of SNORD50A with 68.8% sequence identity, was not detected (Supplementary Table S1). Third, RT-qPCR analyses of our bicistronic reporter showed that SNORD50A overexpression or depletion led to altered mRNA 3′ processing efficiency at the tested PAS. Together, these data strongly suggest that snoRNAs directly regulate mRNA 3′ processing.

Further studies are needed to fully understand the scope, mechanisms, and functional significance of snoRNA or other non-coding RNA-mediated regulation of mRNA 3′ processing. First, it will be important to identify other snoRNAs that also regulate mRNA 3′ processing. In this study, we identified a subset of snoRNAs in the mRNA 3′ processing complex assembled on the SVL PAS. It is possible that distinct snoRNAs or other types of non-coding RNAs may associate with the mRNA 3′ processing com-

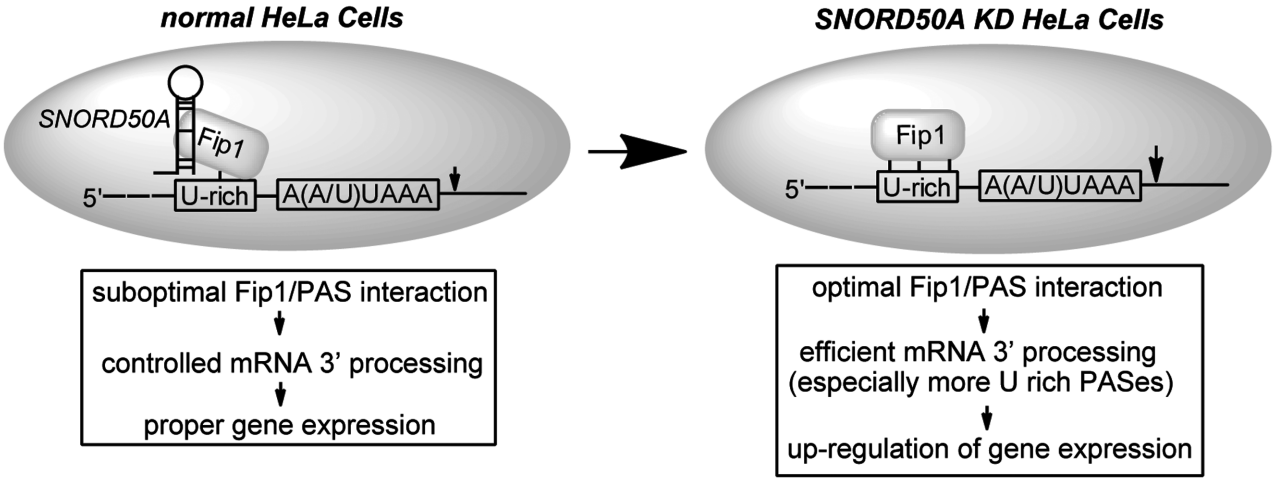


Figure 7. A schematic model of SNORD50A-mediated regulation of mRNA 3′ processing. In normal HeLa cells, SNORD50A is assembled into CPSF via Fip1, and attentuates the interaction of Fip1 with the U rich USE region of PASes, the suboptimal interactions between Fip1 and PASes ensure controlled mRNA 3′ processing activity and proper gene expressions. However, in SNORD50A KD HeLa cells, Fip1 in general tends to bind more frequently to PASes, and thus increased the 3′ processing and gene expression of target mRNAs containing more U rich sequences at the USE of PAS.

plexes and/or regulate mRNA 3′ processing. Indeed, Y3** ncRNA has recently been reported to promote the histone pre-mRNAs 3′ processing by enhancing the recruitment of the CPSF to histone pre-mRNA (61). Second, it will be important to understand how snoRNAs can regulate specific mRNA targets by binding to general mRNA 3′ processing factors. Although the mechanisms are not fully understood, this could be due to the fact that general mRNA 3′ processing factors themselves regulate the mRNA processing and APA of specific mRNAs (8,9,11,20,48,49). Finally, it will be important to understand the functional significance of snoRNA-mediated regulation of mRNA 3′ processing. For example, SNORD50A/B gene locus is lost in 10–40% of 12 common cancers, and the mechanism may involves the direct association of SNORD50A/B with K-Ras (33). Our results showed that SNORD50A-regulated targets are highly enriched for genes that function in proliferation and apoptosis (Supplementary Figure S3A), raising the possibility that snoRNA-mediated mRNA 3′ processing may contribution to cancer development. A key challenge for future studies is the potential functional redundancy of snoRNAs. For example, SNORD50A knockout mice only displayed minor defects (32), raising the possibility that other snoRNAs may at least partially compensate for its loss in its canonical and noncanonical functions.

ACCESSION NUMBERS

All the deep sequencing data have been deposited to GEO database with the accession no. GSE98221.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Dr. Tamas Kiss for providing pCMV-globin plasmid.

FUNDING

National Natural Science Foundation for Youth [31501184]; an interdisciplinary subject grant from the Sun Yat-Sen University and start-up funds from Sun Yat-Sen University (to C.Y.); NIH [GM090056, CA17488]; American Cancer Society [RSG-12–186 to Y.S.]. Funding for open access charge: Start-up grant(to C.Y.).
Conflict of interest statement. None declared.

REFERENCES

1. Colgan,D.F. and Manley,J.L. (1997) Mechanism and regulation of mRNA polyadenylation. *Genes Dev.*, **11**, 2755–2766.
2. Proudfoot,N.J. (2011) Ending the message: poly(A) signals then and now. *Genes Dev.*, **25**, 1770–1782.
3. Gruber,A.R., Martin,G., Keller,W. and Zavolan,M. (2014) Means to an end: mechanisms of alternative polyadenylation of messenger RNA precursors. *Wiley Interdiscip. Rev. RNA*, **5**, 183–196.
4. Shi,Y. and Manley,J.L. (2015) The end of the message: multiple protein-RNA interactions define the mRNA polyadenylation site. *Genes Dev.*, **29**, 889–897.
5. Bienroth,S., Wahle,E., Suter-Crazzolara,C. and Keller,W. (1991) Purification of the cleavage and polyadenylation factor involved in the 3′-processing of messenger RNA precursors. *J. Biol. Chem.*, **266**, 19768–19776.
6. Murthy,K.G. and Manley,J.L. (1992) Characterization of the multisubunit cleavage-polyadenylation specificity factor from calf thymus. *J. Biol. Chem.*, **267**, 14804–14811.
7. Shi,Y., Di Giammartino,D.C., Taylor,D., Sarkeshik,A., Rice,W.J., Yates,J.R. 3rd, Frank,J. and Manley,J.L. (2009) Molecular architecture of the human pre-mRNA 3′ processing complex. *Mol. Cell*, **33**, 365–376.
8. Yao,C., Biesinger,J., Wan,J., Weng,L., Xing,Y., Xie,X. and Shi,Y. (2012) Transcriptome-wide analyses of CstF64-RNA interactions in global regulation of mRNA alternative polyadenylation. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 18773–18778.
9. Yao,C., Choi,E.A., Weng,L., Xie,X., Wan,J., Xing,Y., Moresco,J.J., Tu,P.G., Yates,J.R. 3rd and Shi,Y. (2013) Overlapping and distinct functions of CstF64 and CstF64tau in mammalian mRNA 3′ processing. *RNA*, **19**, 1781–1790.
10. Chan,S.L., Huppertz,I., Yao,C., Weng,L., Moresco,J.J., Yates,J.R. 3rd, Ule,J., Manley,J.L. and Shi,Y. (2014) CPSF30 and Wdr33 directly bind to AAUAAA in mammalian mRNA 3′ processing. *Genes Dev.*, **28**, 2370–2380.

11. Di Giammartino,D.C., Li,W., Ogami,K., Yashinskie,J.J., Hoque,M., Tian,B. and Manley,J.L. (2014) RBBP6 isoforms regulate the human polyadenylation machinery and modulate expression of mRNAs with AU-rich 3′ UTRs. *Genes Dev.*, **28**, 2248–2260.
12. Schonemann,L., Kuhn,U., Martin,G., Schafer,P., Gruber,A.R., Keller,W., Zavolan,M. and Wahle,E. (2014) Reconstitution of CPSF active in polyadenylation: recognition of the polyadenylation signal by WDR33. *Genes Dev.*, **28**, 2381–2393.
13. Di Giammartino,D.C., Nishida,K. and Manley,J.L. (2011) Mechanisms and consequences of alternative polyadenylation. *Mol. Cell*, **43**, 853–866.
14. Shi,Y. (2012) Alternative polyadenylation: new insights from global analyses. *RNA*, **18**, 2105–2117.
15. Elkon,R., Ugalde,A.P. and Agami,R. (2013) Alternative cleavage and polyadenylation: extent, regulation and function. *Nat. Rev. Genet.*, **14**, 496–506.
16. Berkovits,B.D. and Mayr,C. (2015) Alternative 3′ UTRs act as scaffolds to regulate membrane protein localization. *Nature*, **522**, 363–367.
17. Mayr,C. (2016) Evolution and biological roles of alternative 3′UTRs. *Trends Cell Biol.*, **26**, 227–237.
18. Tian,B. and Manley,J.L. (2016) Alternative polyadenylation of mRNA precursors. *Nat. Rev. Mol. Cell. Biol.*, **18**, 18–30.
19. Mayr,C. and Bartel,D.P. (2009) Widespread shortening of 3′UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell*, **138**, 673–684.
20. Lackford,B., Yao,C., Charles,G.M., Weng,L., Zheng,X., Choi,E.A., Xie,X., Wan,J., Xing,Y., Freudenberg,J.M. *et al.* (2014) Fip1 regulates mRNA alternative polyadenylation to promote stem cell self-renewal. *EMBO J.*, **33**, 878–889.
21. Bratkovic,T. and Rogelj,B. (2014) The many faces of small nucleolar RNAs. *Biochim. Biophys. Acta*, **1839**, 438–443.
22. Dupuis-Sandoval,F., Poirier,M. and Scott,M.S. (2015) The emerging landscape of small nucleolar RNAs in cell biology. *Wiley Interdiscipl. Rev. RNA*, **6**, 381–397.
23. Lestrade,L. and Weber,M.J. (2006) snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. *Nucleic Acids Res.*, **34**, D158–D162.
24. Kishore,S. and Stamm,S. (2006) The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C. *Science*, **311**, 230–232.
25. Ender,C., Krek,A., Friedlander,M.R., Beitzinger,M., Weinmann,L., Chen,W., Pfeffer,S., Rajewsky,N. and Meister,G. (2008) A human snoRNA with microRNA-like functions. *Mol. Cell*, **32**, 519–528.
26. Brameier,M., Herwig,A., Reinhardt,R., Walter,L. and Gruber,J. (2011) Human box C/D snoRNAs with miRNA like functions: expanding the range of regulatory RNAs. *Nucleic Acids Res.*, **39**, 675–686.
27. Schubert,T., Pusch,M.C., Diermeier,S., Benes,V., Kremmer,E., Imhof,A. and Langst,G. (2012) Df31 protein and snoRNAs maintain accessible higher-order structures of chromatin. *Mol. Cell*, **48**, 434–444.
28. Falaleeva,M., Pages,A., Matuszek,Z., Hidmi,S., Agranat-Tamir,L., Korotkov,K., Nevo,Y., Eyras,E., Sperling,R. and Stamm,S. (2016) Dual function of C/D box small nucleolar RNAs in rRNA modification and alternative pre-mRNA splicing. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, E1625–E1634.
29. Dong,X.Y., Guo,P., Boyd,J., Sun,X., Li,Q., Zhou,W. and Dong,J.T. (2009) Implication of snoRNA U50 in human breast cancer. *J. Genet. Genomics*, **36**, 447–454.
30. Dong,X.Y., Rodriguez,C., Guo,P., Sun,X., Talbot,J.T., Zhou,W., Petros,J., Li,Q., Vessella,R.L., Kibel,A.S. *et al.* (2008) SnoRNA U50 is a candidate tumor-suppressor gene at 6q14.3 with a mutation associated with clinically significant prostate cancer. *Hum. Mol. Genet.*, **17**, 1031–1042.
31. Pacilli,A., Ceccarelli,C., Trere,D. and Montanaro,L. (2013) SnoRNA U50 levels are regulated by cell proliferation and rRNA transcription. *Int. J. Mol. Sci.*, **14**, 14923–14935.
32. Soeno,Y., Fujita,K., Kudo,T., Asagiri,M., Kakuta,S., Taya,Y., Shimazu,Y., Sato,K., Tanaka-Fujita,R., Kubo,S. *et al.* (2013) Generation of a mouse model with down-regulated U50 snoRNA (SNORD50) expression and its organ-specific phenotypic modulation. *PLoS One*, **8**, e72105.
33. Siprashvili,Z., Webster,D.E., Johnston,D., Shenoy,R.M., Ungewickell,A.J., Bhaduri,A., Flockhart,R., Zarnegar,B.J., Che,Y., Meschi,F. *et al.* (2016) The noncoding RNAs SNORD50A and SNORD50B bind K-Ras and are recurrently deleted in human cancer. *Nat. Genet.*, **48**, 53–58.
34. Kleiman,F.E. and Manley,J.L. (2001) The BARD1-CstF-50 interaction links mRNA 3′ end formation to DNA damage and tumor suppression. *Cell*, **104**, 743–753.
35. Hunt,A.G. (2015) A rapid, simple, and inexpensive method for the preparation of strand-specific RNA-Seq libraries. *Methods Mol. Biol.*, **1255**, 195–207.
36. Mata,J. (2013) Genome-wide mapping of polyadenylation sites in fission yeast reveals widespread alternative polyadenylation. *RNA Biol.*, **10**, 1407–1414.
37. Shepard,P.J., Choi,E.A., Lu,J., Flanagan,L.A., Hertel,K.J. and Shi,Y. (2011) Complex and dynamic landscape of RNA polyadenylation revealed by PAS-Seq. *RNA*, **17**, 761–772.
38. Kaufmann,I., Martin,G., Friedlein,A., Langen,H. and Keller,W. (2004) Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. *EMBO J.*, **23**, 616–626.
39. Keller,W., Bienroth,S., Lang,K.M. and Christofori,G. (1991) Cleavage and polyadenylation factor CPF specifically interacts with the pre-mRNA 3′ processing signal AAUAAA. *EMBO J.*, **10**, 4241–4249.
40. Murthy,K.G. and Manley,J.L. (1995) The 160-kD subunit of human cleavage-polyadenylation specificity factor coordinates pre-mRNA 3′-end formation. *Genes Dev.*, **9**, 2672–2683.
41. Takagaki,Y., Seipelt,R.L., Peterson,M.L. and Manley,J.L. (1996) The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. *Cell*, **87**, 941–952.
42. Lyman,S.K., Gerace,L. and Baserga,S.J. (1999) Human Nop5/Nop58 is a component common to the box C/D small nucleolar ribonucleoproteins. *RNA*, **5**, 1597–1604.
43. Kishore,S., Gruber,A.R., Jedlinski,D.J., Syed,A.P., Jorjani,H. and Zavolan,M. (2013) Insights into snoRNA biogenesis and processing from PAR-CLIP of snoRNA core proteins and small RNA sequencing. *Genome Biol.*, **14**, R45.
44. Darzacq,X., Jady,B.E., Verheggen,C., Kiss,A.M., Bertrand,E. and Kiss,T. (2002) Cajal body-specific small nuclear RNAs: a novel class of 2′-O-methylation and pseudouridylation guide RNAs. *EMBO J.*, **21**, 2746–2756.
45. Marnef,A., Richard,P., Pinzon,N. and Kiss,T. (2014) Targeting vertebrate intron-encoded box C/D 2′-O-methylation guide RNAs into the Cajal body. *Nucleic Acids Res.*, **42**, 6616–6629.
46. Liang,X.H., Vickers,T.A., Guo,S. and Crooke,S.T. (2011) Efficient and specific knockdown of small non-coding RNAs in mammalian cells and in mice. *Nucleic Acids Res.*, **39**, e13.
47. Sharma,E., Sterne-Weiler,T., O'Hanlon,D. and Blencowe,B.J. (2016) Global mapping of human RNA-RNA interactions. *Mol. Cell*, **62**, 618–626.
48. Li,W., You,B., Hoque,M., Zheng,D., Luo,W., Ji,Z., Park,J.Y., Gunderson,S.I., Kalsotra,A., Manley,J.L. *et al.* (2015) Systematic profiling of poly(A)+ transcripts modulated by core 3′ end processing and splicing factors reveals regulatory rules of alternative cleavage and polyadenylation. *PLoS Genet.*, **11**, e1005166.
49. Yang,Y., Li,W., Hoque,M., Hou,L., Shen,S., Tian,B. and Dynlacht,B.D. (2016) PAF complex plays novel subunit-specific roles in alternative cleavage and polyadenylation. *PLoS Genet.*, **12**, e1005794.
50. Konig,J., Zarnack,K., Rot,G., Curk,T., Kayikci,M., Zupan,B., Turner,D.J., Luscombe,N.M. and Ule,J. (2010) iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat. Struct. Mol. Biol.*, **17**, 909–915.
51. Martin,G., Gruber,A.R., Keller,W. and Zavolan,M. (2012) Genome-wide analysis of pre-mRNA 3′ end processing reveals a decisive role of human cleavage factor I in the regulation of 3′ UTR length. *Cell Rep.*, **1**, 753–763.
52. Misra,A. and Green,M.R. (2016) From polyadenylation to splicing: dual role for mRNA 3′ end formation factors. *RNA Biol.*, **13**, 259–264.
53. Moore,C.L. and Sharp,P.A. (1984) Site-specific polyadenylation in a cell-free reaction. *Cell*, **36**, 581–591.

54. Hashimoto,C. and Steitz,J.A. (1986) A small nuclear ribonucleoprotein associates with the AAUAAA polyadenylation signal in vitro. *Cell*, **45**, 581–591.
55. Ryner,L.C. and Manley,J.L. (1987) Requirements for accurate and efficient mRNA 3′ end cleavage and polyadenylation of a simian virus 40 early pre-RNA in vitro. *Mol. Cell. Biol.*, **7**, 495–503.
56. Ellis,J.C., Brown,D.D. and Brown,J.W. (2010) The small nucleolar ribonucleoprotein (snoRNP) database. *RNA*, **16**, 664–666.
57. Fong,Y.W., Ho,J.J., Inouye,C. and Tjian,R. (2014) The dyskerin ribonucleoprotein complex as an OCT4/SOX2 coactivator in embryonic stem cells. *Elife*, **3**, doi:10.7554/eLife.03573.
58. Jinn,S., Brandis,K.A., Ren,A., Chacko,A., Dudley-Rucker,N., Gale,S.E., Sidhu,R., Fujiwara,H., Jiang,H., Olsen,B.N. *et al.* (2015) snoRNA U17 regulates cellular cholesterol trafficking. *Cell Metab.*, **21**, 855–867.
59. Lam,Y.W. and Trinkle-Mulcahy,L. (2015) New insights into nucleolar structure and function. *F1000Prime Rep.*, **7**, 48.
60. Youssef,O.A., Safran,S.A., Nakamura,T., Nix,D.A., Hotamisligil,G.S. and Bass,B.L. (2015) Potential role for snoRNAs in PKR activation during metabolic stress. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 5023–5028.
61. Kohn,M., Ihling,C., Sinz,A., Krohn,K. and Huttelmaier,S. (2015) The Y3** ncRNA promotes the 3′ end processing of histone mRNAs. *Genes Dev.*, **29**, 1998–2003.